AP Stat- Unit 1 Review (Chapters 2 - 6) Key

Multiple Choice Questions

1. A random sample of 25 birthweights (in ounces) is taken yielding the following summary statistics:

Variable	N	Mean	Median	~~TrMean~~	StDev	~~SE Mean~~
Birthwt	25	129.40	129.00	~~128.35~~	17.41	~~3.48~~
Variable	Minimum	Maximum	Q1	Q3		
Birthwt	96.00	187.00	119.50	135.50		

What can be said about the number of outliers for this data set?

(A) 0
(B) At least 1
(C) No more than 1
(D) At least 2
(E) No more than 2

UF = 159.5 = (1.5×IQR) + Q3
LF = 95.5 = Q1 − (1.5×IQR)

B

2. For a set of values, suppose the mean is 10 and the standard deviation is 2. If each value is multiplied by 9 and added by 10, what will be the mean and standard deviation for this new set of values?

~~(A)~~ mean 10; standard deviation 2
~~(B)~~ mean 10; standard deviation 18
~~(C)~~ mean 100; standard deviation 2
(D) mean 100; standard deviation 18
(E) mean 100; standard deviation 28

x9 + 10

D

3. In this year's county mathematics competition, a student scored 40; in last year's competition, the student scored 35. The average score this year was 38 with a standard deviation of 2. Last year's average score was 34 with a standard deviation of 1. In which year did the student score better?

(A) The student scored better on this year's exam
(B) The student scored better on last year's exam
(C) The student scored equally well on both exams
(D) Without knowing the number of test items, it is impossible to determine the better score.
(E) Without knowing the number of students taking the exam in the county, it is impossible to determine the better score.

This: 40 $\mu = 38$ $\sigma = 2$
Last: 35 $\mu = 34$ $\sigma = 1$

This $z = 1$
Last $z = 1$ } same

C

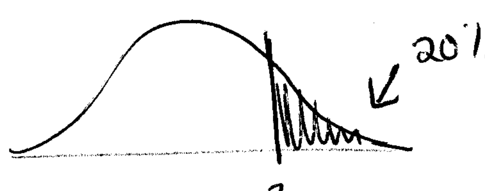
4. The heights of American men aged 18 to 24 are approximately normal with a mean of 68 inches and a standard deviation of 2.5 inches. About 20% of these men are taller than

(A) 66 inches
(B) 68 inches
(C) 70 inches
(D) 72 inches
(E) 74 inches

higher

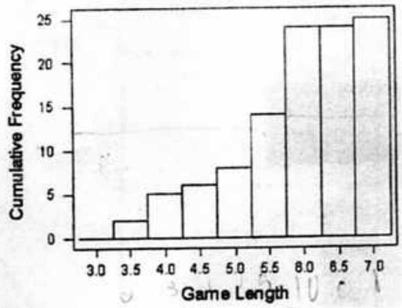
N(68, 2.5)

C



invnorm(0.80, 68, 2.5)

5. The lengths (in innings) of 25 randomly selected Little League baseball games were recorded, and a *cumulative* frequency histogram was created from the results. What is the best conclusion that can be made from the graph?

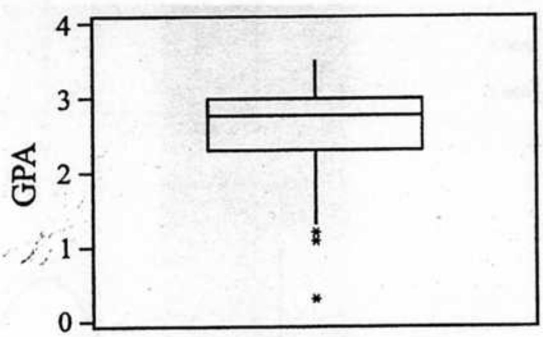


total = 25 observations
$M = 12.5^{th}$ observation

D

(A) ~~The median game length is 5 innings~~. $M = 5.5$
(B) ~~Fourteen games lasted 5.5 innings~~ or less.
(C) ~~A majority of the games lasted 6 or more innings~~
(D) The distribution of game lengths is severely skewed left
(E) Games lasting more than 6 innings occurred least frequently

6. Which statement is true about the boxplot below?



I. It is a left skewed distribution which has outliers ✓
~~II.~~ It is a symmetrical distribution which has outliers
III. The interquartile range is less than 1 ✓
IV. Approximately 75% of the observations have a GPA less than 3 ✓

E

(A) I only
(B) II only
(C) II and III
(D) III and IV only
(E) I, III, and IV

7. The scores of a standardized test designed to measure math anxiety are normally distributed with a mean of 100 and a standard deviation of 10 for a population of first year college students. Which of the following observations would you suspect is an outlier?

C

(A) 90
(B) 100
(C) 150
(D) 90, 100, and 150 are all outliers
(E) None of 90, 100, and 150 are outliers

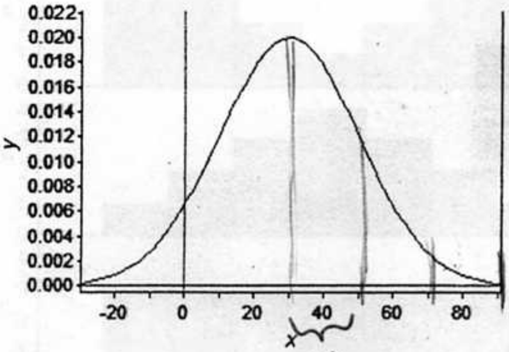
$N(\overset{\mu}{100}, \overset{\sigma}{10})$

$\mu \pm 2\sigma = (80, 120)$

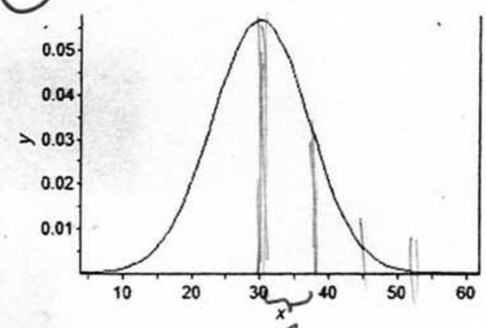
8. Which of the following distributions has a mean of 30 and a standard deviation of 7?

B

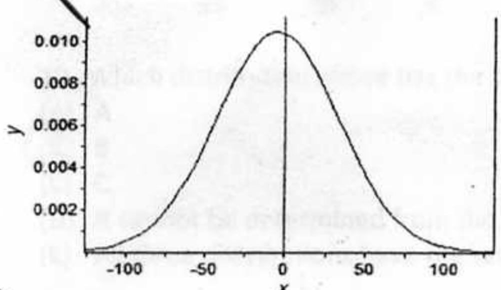
(A)



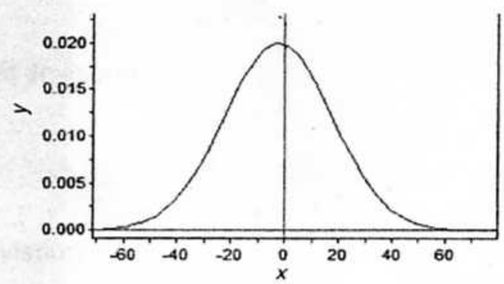
(B)



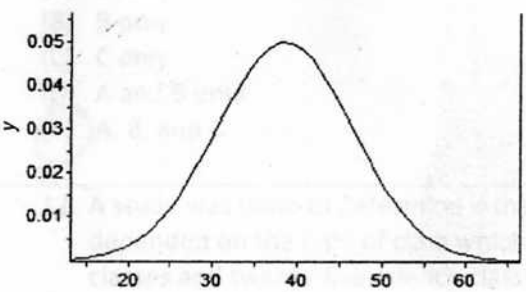
~~(C)~~



~~(D)~~



~~(E)~~



9. A researcher interested in the age at which women are having their first child surveyed a simple random sample of 250 women having at least one child and found a approximately normal distribution with a mean age of 27.3 and a standard deviation of 5.4. Approximately 95% of the women had their first child between the ages of

B

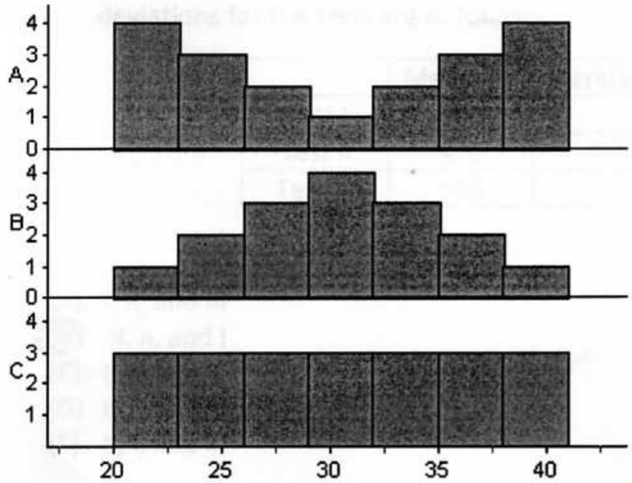
(A) 11.1 years and 43.5 years
(B) 16.5 years and 38.1 years
(C) 21.9 years and 32.7 years
(D) 21.9 years and 38.1 years
(E) 25.0 years and 29.6 years

$\mu \quad \sigma$

$N(27.3, 5.4)$

$\mu \pm 2\sigma =$

Use the following for questions 10 and 11.



B

10. Which distribution above has the smallest standard deviation?
(A) A
(B) B
(C) C
(D) It cannot be determined from the graphs
(E) All three distributions have the same standard deviation

E

11. In which distribution(s) would you be more likely to find the mean and median the same?
(A) A only
(B) B only
(C) C only
(D) A and B only
(E) A, B, and C

---

12. A study was done to determine if the method of instruction (either lecture or discussion) depended on the type of class which was being taught. Twenty art classes, seventeen math classes and twenty-five science classes were observed. The method of instruction, discussion or lecture, was recorded. Which of the following best describes the relationship between method of teaching and type of class?

C

	Discuss	Lecture	
Arts	5	15	20
Math	12	5	17
Science	15	10	25
	32	30	62

(A) There appears to be no relationship since the number of discussion class and the number of lecture classes was exactly the same
(B) No association can be determined since the number of art, math, and science classes were not exactly the same
(C) There appears to be an association since the art class was less likely to use discussion than either math or science
(D) There appears to be an association since the number of math and science classes is greater than the number of arts classes
(E) A measure of association cannot be determined from these data

13. Polly takes three standardized tests. She scores 600 on all three. Using standard scores, or z-scores, rank her performance on the three tests from best to worst if the means and standard deviations for the tests are as follows:

	Mean	Standard Deviation	
Test I	500	80	$z = 1.25$
Test II	470	120	$z = 1.083$
Test III	560	30	$z = 1.33$

B

(A) I, II, and III
(B) III, II, and I
(C) I, III, and II
(D) III, I, and II
(E) II, I, and III

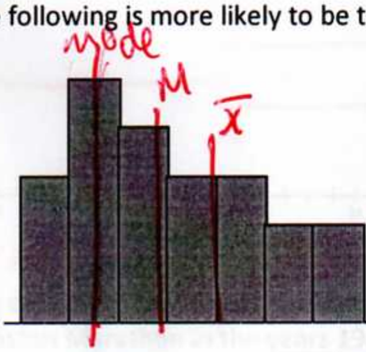
14. Which of the following will most likely approximate a uniform distribution?

E

(A) Heights of students at a particular high school
(B) Weights of students at a particular high school
(C) SAT scores of seniors at a particular high school
(D) IQ scores of students at a particular high school
(E) Ages of students at a particular high school

normal shape

15. Which of the following is more likely to be true of this distribution?



mode < Med < mean

C

(A) Mean = 3	Median = 3	Mode = 3
(B) Mean = 3.5	Median = 4	Mode = 3
(C) Mean = 4	Median = 3.5	Mode = 3
(D) Mean = 3.5	Median = 3.5	Mode = 5
(E) Mean = 3	Median = 2	Mode = 5

16. If the standard deviation of a distribution is 4, the variance is:

D

(A) 4
(B) 2
(C) 8
(D) 16
(E) 0

s^2 = variance

Free Response

1. Make a back-to-back split stemplot of the following data: → over →

Reading Scores

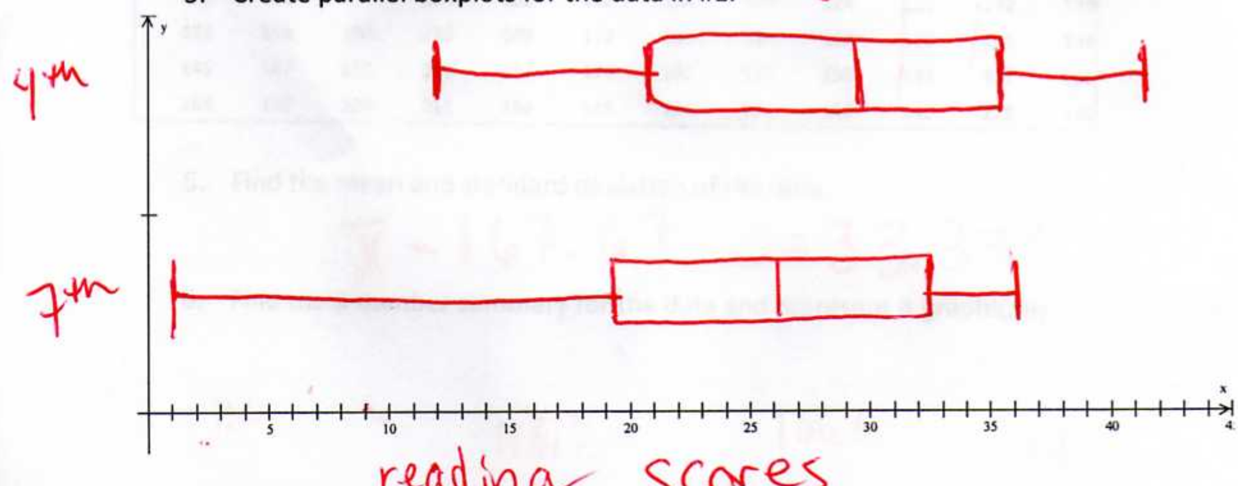
4^{th} Graders	12	15	18	20	20	22	25	26	28	29
	31	32	35	35	35	36	37	39	40	42
7^{th} Graders	1	12	15	18	18	20	23	23	24	25
	27	28	30	30	31	33	33	33	35	36

shape
center
spread

2. Make a comparison between 4^{th} grade and 7^{th} grade reading scores based on your stemplot.

4th graders are roughly symmetric, but 7th grade are left skewed. 4th graders have a center @ the mean of 28.85 and 7th graders have a center @ the median of 26. The range of 4th graders is from 12 – 42. the range of (1, 36) is for 7th graders

3. Create parallel boxplots for the data in #1.

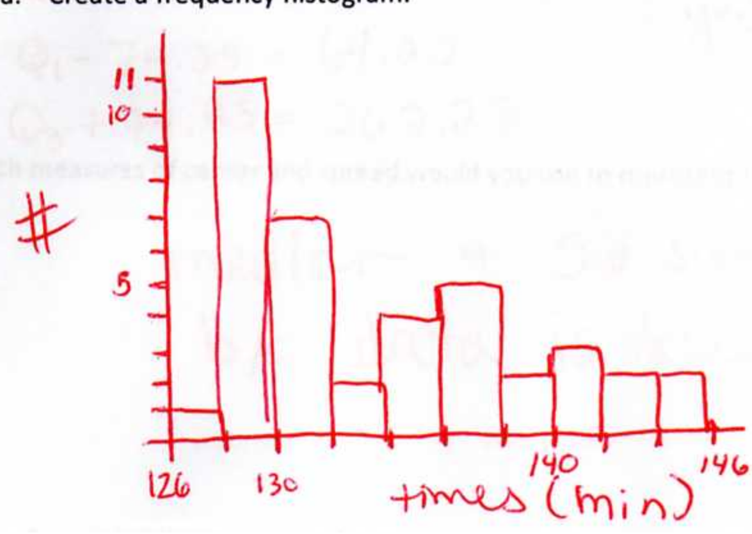


4. The following data gives the times (in minutes, rounded to the nearest minute) for the winning man in the Boston Marathon in the years 1959 to 1997.

Times:

143	139	136	139	130	129	129	132	131	129
141	140	142	136	140	132	131	129	128	129
144	137	134	136	135	129	134	129	130	131
144	137	131	134	130	129	128	128	127	

a. Create a frequency histogram.



1

4th		7th
	0	1
	0	
2	1	2
8 5	1	5 8 8
2 0 0	2	0 3 3 4
9 8 6 5	2	5 7 8
2 1	3	0 0 1 3 3 3
9 7 6 5 5 5	3	5 6
2 0	4	
	4	

b. How would a cumulative frequency histogram differ?

you add up the total observations up to that point + put in the bar. Always increasing picture.

c. Describe the histogram with appropriate summary statistics.

* don't say mean b/c it is skewed!

The distribution is right skewed. So the center is the median of 132 and the spread is (127, 144), with an IQR of 8 units.

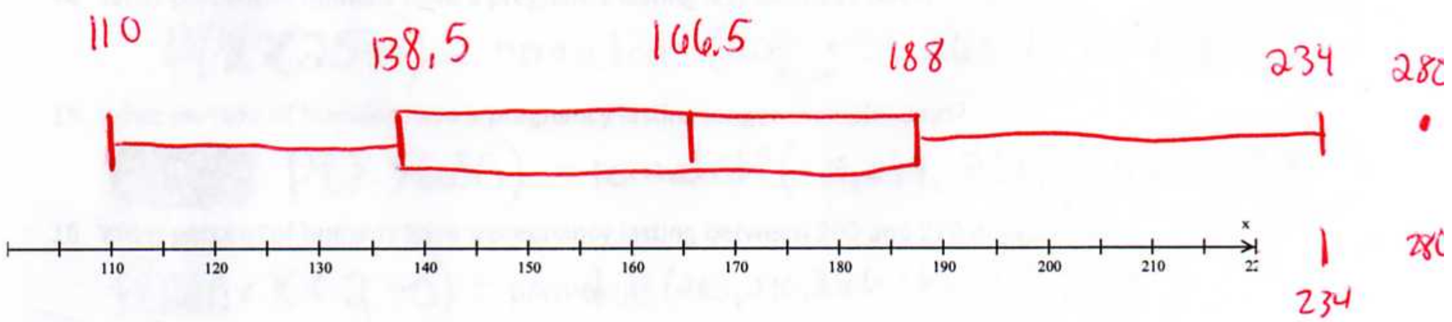
Use the following set of data for numbers 5-8.

weights of 18 year old males											
130	201	190	234	188	162	134	120	124	199	142	179
128	158	202	280	189	172	135	110	187	200	165	188
145	167	122	151	187	174	162	132	195	132	137	186
188	166	204	211	164	145	184	124	160	140	175	180

5. Find the mean and standard deviation of the data.

$\bar{x} = 167.67 \quad s = 33.378$

6. Find the 5 number summary for the data and represent it graphically. (boxplot)



7. Test for possible outliers. Are there any in your opinion (based on the data and the test)?

$IQR = 49.5 \times 1.5 = 74.25$

$Q_1 - 74.25 = 64.25$

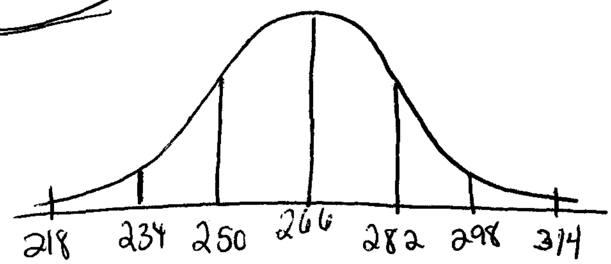
$Q_3 + 74.25 = 262.25$

yes - 280

8. Which measures of center and spread would you use to represent this data? Why?

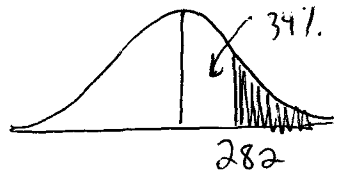
median + 5# summary b/c data is skewed.

9. Sketch the graph of N(266, 16), the distribution of pregnancy length from conception to birth for humans.



Use the empirical rule (the 68-95-99.7 rule) for problems 10-12.

10. Find the length of the longest 16% of all pregnancies. Sketch and shade a normal curve for this situation.



282 days

11. Find the length of the middle 99.7% of all pregnancies.

(218, 314) days

12. Find the length of the shortest 2.5% of all pregnancies.

234 days

13. What z-score does a pregnancy of 257 days have?

$z = \frac{257-266}{16} = 0.0604$

14. What percent of humans have a pregnancy lasting less than 257 days?

$P(X < 257) =$ normalcdf(257, E99, 266, 16) = −0.2869

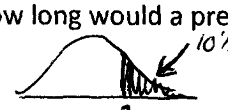
15. What percent of humans have a pregnancy lasting longer than 280 days?

$P(X > 280)$ = normalcdf(280, E99, 266, 16) = 0.1908

16. What percent of humans have a pregnancy lasting between 260 and 270 days?

$P(260 < X < 270)$ = normalcdf(260, 270, 266, 16) = 0.2449

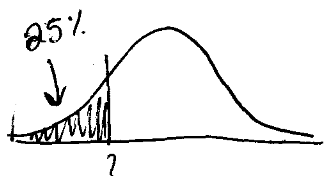
17. How long would a pregnancy have to last to be in the longest 10% of all pregnancies?



$P(X > ?) = 10\%$ or $P(X < ?) = 90\%$ invnorm(0.90, 266, 16)

= $\boxed{286.5048 \text{ days}}$

18. How short would a pregnancy be to be in the shortest 25% of all pregnancies?



$P(X < ?) = 0.25$

invnorm(0.25, 266, 16) =

$\boxed{255.2082 \text{ days}}$

19. The life expectancy of a particular brand of light bulb is normally distributed with a mean of 1500 hours and a standard deviation of 75 hours. $N(1500, 75)$

a. What is the probability that a light bulb will last less than 1410 hours?

$P(x < 1410) = \text{normalcdf}(-E99, 1410, 1500, 75) = 0.1151$

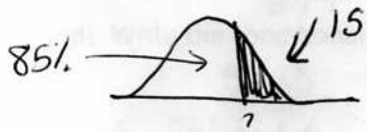
b. What is the probability that a light bulb will last more than 1550 hours?

$P(x > 1550) = \text{normalcdf}(1550, E99, 1500, 75) = 0.2525$

c. What is the probability that a light bulb will last between 1563 and 1648 hours?

$P(1563 < X < 1648) = \text{normalcdf}(1563, 1648, 1500, 75) = 0.1762$

d. 15% of the time a light bulb will last more than how many hours?

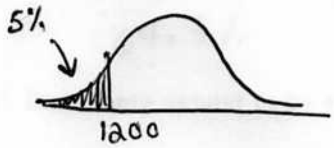


$P(x < ?) = 0.85$ invnorm(0.85, 1500, 75)

$\boxed{= 1577.73 \text{ hours}}$

e. If we wanted only 5% of the bulbs to last less than 1200 hours and we can't change the standard deviation, what must the new mean hours become?

$N(\mu, 75)$

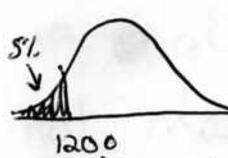


$z = \text{invnorm}(0.05, 0, 1) = -1.645$

$-1.645 = \frac{1200 - \mu}{75}$

$\boxed{\mu = 1323.375 \text{ hours}}$

f. If we wanted only 5% of the bulbs to last less than 1200 hours and we can't change the mean hours, what must the new standard deviation become? $N(1500, \sigma)$



$z = \text{invnorm}(0.05, 0, 1) = -1.645$

$-1.645 = \frac{1200 - 1500}{\sigma}$

$\boxed{\sigma = 182.371 \text{ hrs}}$

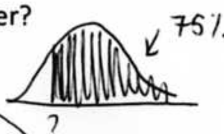
20. A water fountain is designed to dispense a volume of 12.2 oz. with a standard deviation of 0.5 oz.

$\mu = 12.2$ $\sigma = 0.5$ assume normal

a. What percentage of cups end up with at least 12 oz.?

$P(x > 12) = \text{normalcdf}(12, E99, 12.2, 0.5) = 0.6554$

b. 75% of the cups contain more than how much water?

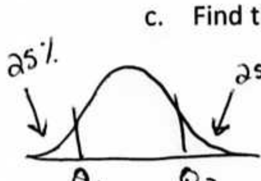


$P(x < ?) = 0.25$

invnorm(0.25, 12.2, 0.5)

$\boxed{= 11.863 \text{ oz}}$

c. Find the IQR for the amount of water dispensed.



$P(x < Q_1) = 0.25$

$P(x < Q_3) = 0.75$

$Q_1 = \text{invnorm}(0.25, 12.2, 0.5) = 11.863 \text{ oz}$

$Q_3 = \text{invnorm}(0.75, 12.2, 0.5) = 12.537 \text{ oz}$

$\boxed{IQR = 0.674 \text{ oz}}$

21. The principal of a school with 484 students collected information about how many of the students wear glasses.

	Always wear glasses	Sometimes wears glasses	Never wear glasses	
Boys	40	48	161	249
Girls	36	55	144	235
	76	103	305	484

(a) Fill in the missing value

(b) Find the marginal distribution of glasses

Always : 15.7%
Sometimes : 21.3%
never : 63.02%

(c) What percent of boys never wear glasses?

$\frac{161}{249} = 64.66\%$

(d) Write the conditional distribution of sex

Boys
A 16.1%
S 19.3%
N 64.7%

Girls
A 15.3%
S 23.4%
N 61.3%

(e) Does there appear to be an association between sex and whether they wear glasses or not?

Although the marginal + conditional distributions calculated above (marginal of glasses and conditional of sex) do not match exactly, they are within 2% of each other, so I would say that the two variables ARE INDEPENDENT. (there is no association btw. sex + glasse wearing).